REMARKS

Claims 1-20 are pending in the application. The IDS was objected to due to the omission of the names of authors and the omission of a reference translation. The abstract was objected to because it exceeded the specified length and referenced other material. Claim 10 was rejected under 35 U.S.C. § 101 as being directed toward non-statutory subject matter. Claims 1, 3, 5, 6, 10-12, 16, and 17 were rejected under 35 U.S.C. § 102 as being unpatentable over U.S. Patent No. 6,104,989 at Kanevsky *et al.* ("Kanevsky"). Claims 4 and 15 were rejected under 35 U.S.C. § 103 as being unpatentable over Kanevsky. Claims 2, 9, 13, and 20 were rejected under 35 U.S.C. § 103 as being unpatentable over Kanevsky in view of NPL document "A tutorial in Hidden Markov Models and Selected Application in Speech Recognition" ("Rabiner"). Finally, claims 7-8 and 18-19 were rejected under 35 U.S.C. § 103 as being unpatentable over Kanevsky in view of NPL document "Bayesian Adaptive Learning of the Parameters Hidden Markov Model for Speech Recognition ("Huo"). Claim 10 has been amended and claims 21-23 have been added.

The IDS was objected to due to the omission of the names of authors and the omission of a reference translation. In response to this objection, Applicant will be filing a Supplemental IDS in due course.

The abstract was objected to because it exceeded the specified length and referenced other material. Applicant has included an amended abstract which complied with the length requirements addressed in the Office Action.

Claim 10 was rejected under 35 U.S.C. § 101 as being directed toward non-statutory subject matter. To clarify the subject matter of claim 10, Applicant has amended the preamble of claim 10 to recite "A computer readable storage medium which stores a computer executable program that when executed by a computer

processor causes a computer to execute the steps of," instead of "A program for causing a computer to execute the steps of."

Claims 1, 3, 5, 6, 10-12, 16, and 17 were rejected under 35 U.S.C. § 102 as being unpatentable over Kanevsky. Claims 1 and 10-12 are independent claims.

Kanevsky discloses a so-called model-based segmentation technique which sequentially applied previously-learned topics (models) to input texts. Overall, Kanevsky's approach appears virtually identical to the second conventional technique disclosed in the specification. [0006]-[0007], [00010].

Claims 1, 10, 11, and 12 recite, *inter alia*, "generating a probability model in which information indicating each word of a text document belongs to a topic is made to correspond to a latent variable." Page 4 of the Office Action alleges that Kanevsky discloses generating a probability model in which information indicating each word of a text document is made to correspond to a latent variable, as required by claims 1, 10, 11, and 12. Applicants respectfully disagree.

Unlike claims 1, 10, 11, and 12, Kanevsky discloses using training data to create a battery of topics for assignment to text strings and ratios for determining changes in topic. *See* col. 2, lns. 39-43. Kanevsky's topic identification method "relies upon a battery of topics formed from existing training data, and uses estimates of probabilities for [a given] segment with respect to the distributions of words from the text for each topic probability model in the battery, and compares estimated scores." Col. 3, lns. 46-51. In contrast, claims 1, 10, 11, and 12 recite no such learning step of using training data nor do they require any training data at all.

Unlike claims 1, 10, 11, and 12, Kanevsky does not create a model "model in which information indicating each word of a text document belongs to a topic is made to correspond to a latent variable." Rather, Kanevsky applies the models it created using

training data, which as previously mentioned, claims 1, 10, 11, and 12 do not utilize. Applying these models, Kanevsky will look at a string of text and assign it a topic. *See* col. 4, lns. 26-28. After a single topic has been assigned to a string of text, Kanevsky will cease accumulating this text in a buffer 402 and will send the segment from the buffer 402 to a machine 405, which will perform translation on this homogenous segment. Col. 9, lns. 1-6, Fig. 4. The machine 405 uses the models trained (on training data) for each topic to translate the text. *See* col. 9, lns. 7-8, Fig. 4.

In this manner, Kanevsky calculates a likelihood of input text corresponding to a topic by using a model of previously-learned topics to determine matching between input text and the topic. In contrast, claims 1, 10, 11, and 12, include a step of generating a probability model, not merely using existing models to analyze input text as taught by Kanevsky.

Claims 1, 10, 11, and 12 recite, *inter alia*, "outputting an initial value of a model parameter which defines the generated probability model." Pages 4-5 of the Office Action allege the Kanevsky discloses this limitation. Applicant respectfully disagrees.

As previously discussed, Kanevsky does not generate a probability model that meets the limitations of claims 1, 10, 11, and 12, *e.g.*, a "model in which information indicating each word of a text document belongs to a topic is made to correspond to a latent variable." Claims 1, 10, 11, and 12 require that the initial value of a model parameter defines the generated probability model. Thus, without a comparable probability model, this limitation also clearly not taught by Kanevsky.

Furthermore, Page 5 of the Office Action compares the Kanevsky's designation of a text string as belonging to a neutral topic to the probability model in claims 1, 10, 11, and 12. Meanwhile, Kanevsky's neutral topic refers to a model of

expediential topics averaged from all topics. The model remains unchanged in text segmentation and is only a placeholder used to restart topic identification based on the battery topics when no topic can be identified. *See* col. 2, lns. 43-44. When a text string does not match any of the topics that were created in the battery by training data, this "neutral topic" is assigned. Contrary to the Office Action's assertions on Page 5, assertions, the neutral topic does not establish a generality that can be further defined. Rather, the neutral topic in Kanevsky segregates the text string to which it is assigned from that being translated unless more training data is used to add an additional topic to the battery. *See* col. 3, lns. 54-55. Meanwhile, the probability model itself in claims 1, 10, 11, and 12 is used to segment the text document.

For at least the foregoing reasons, claims 1, 10, 11, and 12 are patentable over Kanevsky. By virtue of their dependence on one of the listed independent claims, claims 3, 5, 6, 16, and 17 are also patentable over Kanevsky.

Claims 4 and 15 were rejected under 35 U.S.C. § 103 as being unpatentable over Kanevsky. Claim 4 is a dependent claim which depends from claim 1 and claim 15 is a dependent claim which depends from claim 12. As previously discussed, claims 1 and 12 are patentable over Kanevsky. Thus, claims 4 and 15 are patentable over Kanevsky at least by virtue of their dependence on claims 1 and 12.

Finally, claims 2, 9, 13, and 20 were rejected under 35 U.S.C. § 103 as being unpatentable over Kanevsky in view Rabiner. These claims are dependent claims, which depend either from claim 1 or claim 12. As previously discussed, claims 1 and 12 are patentable over Kanevsky. Thus, claims 2, 9, 13, and 20 are patentable over Kanevsky at least by virtue of their dependence on claims 1 and 12 because Rabiner does not supplement the deficiencies in Kanevsky.

Claim 2 discloses, *inter alia*, "selecting a probability model from the plurality of probability models, which is used to perform processing in the step of segmenting the text, on the basis of the plurality of estimated model parameters." Page 9 of the Office Action concedes that Kanevsky does not perform this step, but alleges that because Rabiner discloses multiple HMM models, it would be obvious to develop multiple models for Kanevsky and upon use, to select a model for segmentation of text. Applicant respectfully disagrees with this assertion.

Kanevsky generates probability models for each topic before text is input for segmentation. *See* col. 3, lns. 46-51. Because these probability models represent the fixed topics in the battery they are only used when Kanevsky assigns a topic to a text string. The text is later segmented in accordance with the topic assigned, not in accordance with the probability models, as required by claim 2. Furthermore, the step of segmenting the text is not performed based on a plurality of estimated model parameters, as required by claim 2. Rather, before the text can be segmented, a definite topic from the battery must be assigned to the text and as discussed earlier, these topics were formed based on training data, which the claimed invention does not employ.

Kanevsky does not employ a model in text segmentation. Combining Rabiner with Kanevsky for the purpose of having multiple models will not teach the above limitation of claim 2. Kanevsky and Rabiner combined would still lack the step of selecting a model on the basis of estimated model parameters.

For at least this reason, claim 2 is patentable over Kanevsky in view Rabiner. Claim 9 depends from claim 2 and is therefore patentable over Kanevsky in view Rabiner by virtue of its dependence.

Claim 13 is an apparatus claim that parallels claim 2. Claim 13 recites, *inter alia*, "selecting a probability model from the plurality of probability models." As discussed previously in regards to claim 2, claim 13 is patentable over Kanevsky in view Rabiner for at least the same reasons. As claim 20 depends from claim 13, it is therefore patentable over Kanevsky in view Rabiner by virtue of its dependence.

Claims 7-8 and 18-19 were rejected under 35 U.S.C. § 103 as being unpatentable over Kanevsky in view of Huo. These are all dependent claims that depend from either independent claim 1 or independent claim 12. As previously discussed, claims 1 and 12 are patentable over Kanevsky.

Huo discusses the adaptive training of the parameters in Markov models. Abstract. However, Huo does not disclose the steps of either independent claim 1 or claim 12, which are missing from Kanevsky. Therefore, Huo does not remedy the deficiencies of Kanevsky.

For at least the foregoing reasons, claims 7-8 and 18-19 are patentable over Kanevsky in view of Huo.

Claims 21-23 have been added to more fully recite the present invention. Support for these claims can be found in the specification at page 14, line 27, through page 17, line 14, and from page 19, lines 3 through 21.

In view of the above amendment, applicant believes the pending application is in condition for allowance.

Dated: August 25, 2008

Respectfully submitted,

/Michael J. Scheer/

By_____

Michael J. Scheer
Registration No.: 34,425
DICKSTEIN SHAPIRO LLP
1177 Avenue of the Americas
New York, New York 10036-2714
(212) 277-6500
Attorney for Applicant